THE VALIDITY OF CERTAIN ACCEPTED CRITERIA OF DESIGN

BY

LLEWELLYN NOEL WILEY

B.S., Bradley Polytechnic Institute, 1931

M.S., Purdue University, 1934

AN ABSTRACT OF A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN PSYCHOLOGY
IN THE GRADUATE SCHOOL OF THE
UNIVERSITY OF ILLINOIS, 1940

URBANA, ILLINOIS
1940

THE VALIDITY OF CERTAIN ACCEPTED CRITERIA OF DESIGN

BY

LLEWELLYN NOEL WILEY

B.S., Bradley Polytechnic Institute, 1931

M.S., Purdue University, 1934

AN ABSTRACT OF A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN PSYCHOLOGY
IN THE GRADUATE SCHOOL OF THE
UNIVERSITY OF ILLINOIS, 1940

URBANA, ILLINOIS
1940

ACKNOWLEDGMENTS

The writer is gratefully indebted to Professor Herbert Woodrow for his helpful criticisms and general direction of this study, to the members of the staff of the Psychology Department for their time and advice, and to those on the staff of the School of Fine Arts who acted as subjects. In addition, he particularly wishes to thank those students who gave liberally of their time without material reward.

THE VALIDITY OF CERTAIN ACCEPTED CRITERIA OF DESIGN

STATEMENT OF THE PROBLEM

The purpose of this investigation was to discover what determines that an abstract design will be pleasing or displeasing. The method employed involved factor analyzing "pleasantness" ratings assigned to 50 abstract designs. This resulted in the additional problem of identifying the obtained factors. Thus the study sought to identify the factors involved in rating abstract designs for "pleasantness"; and to do this with the least possible ambiguity.

In order to have criteria with which to compare the obtained factors, it was necessary to have the figures rated for other qualities than "pleasantness". The ones chosen were balance, symmetry, rhythm, and unity, as rated by artists.

In another study in which criteria were sought, students rated upon the above four qualities, and also the properties of "complexity", "meaningfulness", and "depth". Finally, the search for criteria with which to associate the factors of "pleasantness" led to rating the designs upon a quality here captioned as "ease of seeing organization". This perceptual characteristic was described to the students as the degree to which some "lawfulness" of relation among the compositional elements of the pattern can be inferred.

EXPERIMENTAL METHODS

The materials and procedures listed in this section of the abstract cover, in the main, the techniques used throughout several related studies. Two of these separate experiments have been consolidated in this summary under the title "The Rating Scale Experiment and Factor Validation Study". Two preliminary investigations have been omitted, while two others are discussed briefly. Consequently, the procedures outlined in this section of the abstract apply to its three major divisions, namely: (I) "The Rating Scale Experiment and Factor Validation Study"; (II) "The Artists' Rating Experiment"; (III) "The Training Experi-

ment". Materials and techniques mentioned only under these headings are limited to the experiment they accompany.

Materials: Abstract patterns were made into 50 lantern slides. These are shown in Plates I and II. The subjects rated these slides upon scales

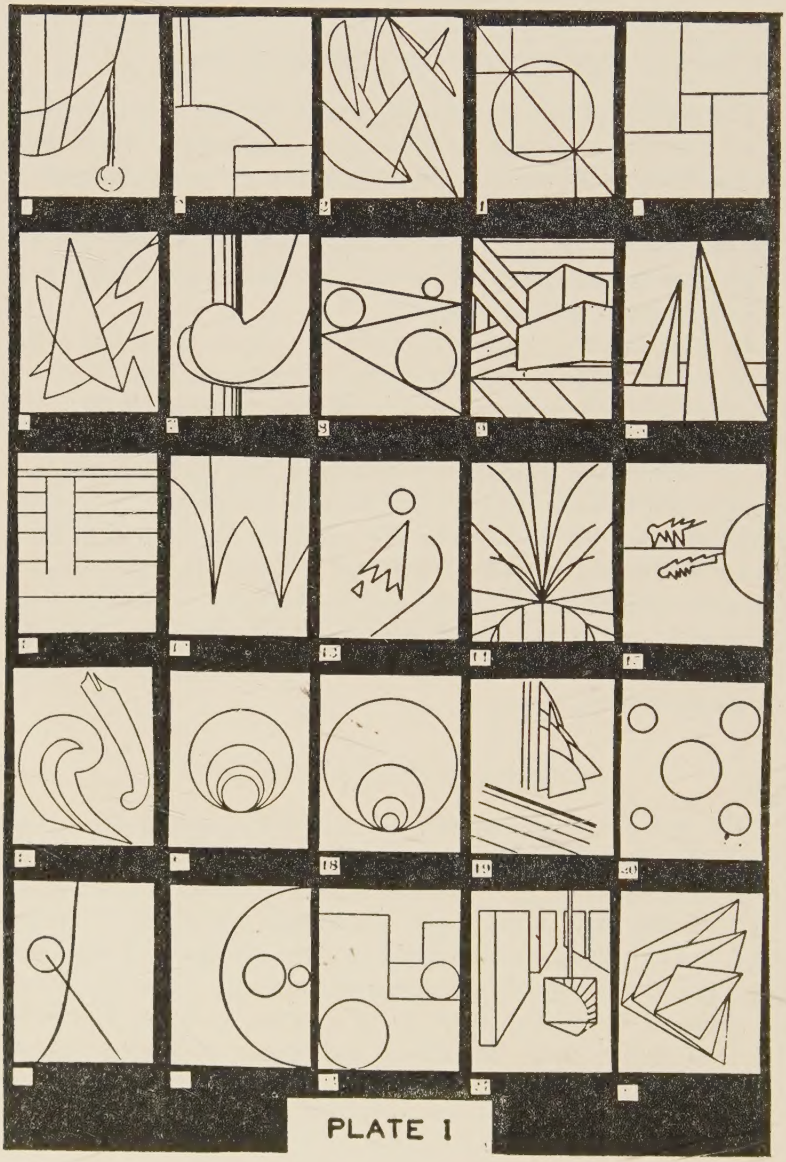

PLATE I

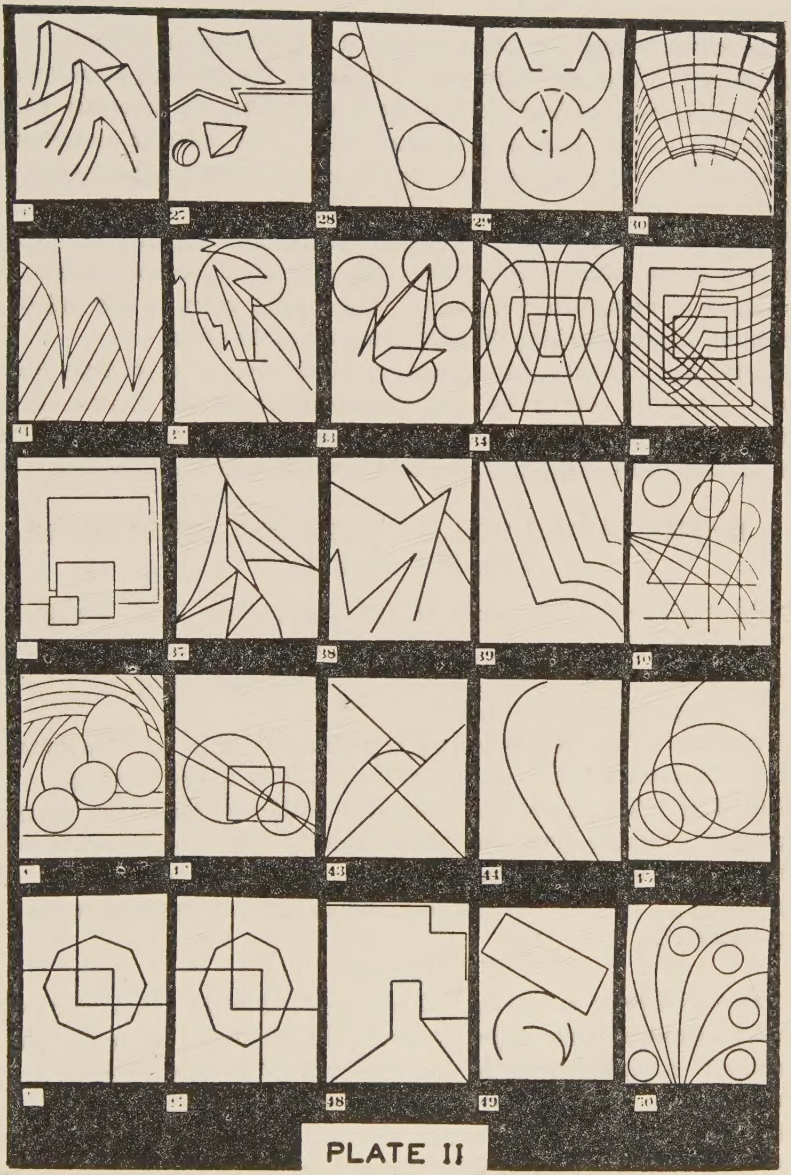
PLATE II

which were divided into 5 segments, each of which was given a descriptive caption.

Subject personnel: Both men and women were used as subjects. All were either students or instructors at the University of Illinois.

Procedure: The procedures varied little for the different experiments. The slides were projected before the subject group for intervals which lasted 15 seconds in most projection series. The order of presentation was random from day to day, but the same sequence was repeated on those few occasions when two ratings followed each other immediately. While the slide was visible the subject rated it for the particular quality named in the rating scale captions. He was told to regard the scale as continuous, despite its 5 divisions; thus a check mark at one end of a segment meant something different from one at the other.

Scoring and calculating correlations: The scoring of most of the scales was upon a 10 point basis, each segment being divided in two. From these scores, singly and pooled, product-moment correlation coefficients were computed as needed. Only the product-moment coefficient was used.

I "THE RATING SCALE EXPERIMENT AND FACTOR VALIDATION STUDY"

Subject personnel: A group of 138 students rated the designs on "pleasantness". Another group of 12 students rated them upon "ease of seeing organization". Not more than 20 of the larger group were trained in art; the others had no training.

Materials: The "ease of seeing organization" scale was the same as the "pleasantness" scale of 5 segments, but with a new heading.

Procedure: The series was presented to 5 different classes for the purpose of obtaining scores on "pleasantness". The two presentations in this portion of the experiment were made two weeks apart. Only one presentation was made to the group which rated on "ease of seeing organization".

Results

The reliability of ratings on "pleasantness". For the 138 subjects who were present on both projection series, the correlation of mean "pleasantness" ratings between the odd and even halves of the subject group was +.978, for the first projection series. For the second series the corresponding coefficient was +.973. Using the Spearman-Brown prophecy formula for 2N, these "stepped up" to +.989, and +.986, re-

spectively. The correlation between the first and second presentation series, based upon mean "pleasantness" ratings, is +.941. So far as group ratings on "pleasantness" are involved, it seems that this quality can be reliably evaluated for this set of designs.

The reliability of ratings on "ease of seeing organization." The correlation between 5 men and one woman on the one hand, and 5 men and a woman on the other, gave a reliability of +.872 for "ease of seeing organization". This "stepped up" to +.932 by using the prophecy formula for 2N. One may conclude that group ratings on this quality are even more reliable than similar ratings on "pleasantness", if groups of equal size are compared.

Individual agreement and stability on "pleasantness" ratings. The "pleasantness" ratings of the individual subjects were correlated for the two presentation series. The resulting distribution of coefficients was approximately normal. Its mean was +.50.

Samples of 50 pairs of subjects were chosen at random from the 138. The mean of their interperson correlations was about +.25 for the first presentation series, and +.40 for the second.

It may be concluded from the above correlations that the individual agrees with himself to the extent of an average retest correlation of +.50 when rating these figures on "pleasantness". Agreement between persons is not quite as strong.

The choice of "tests" and ratings to be factor analyzed. It was possible to select 50 persons whose retest correlations exceeded +.57. The two ratings these subjects gave every pattern were summed by individuals. These sums were intercorrelated for 14 selected designs. The choice of the 14 tests was made from data obtained in the next experiment to be described, i.e., "The Artists' Rating Experiment". The criterion for the selection was an uneven rating of the design in some one of the four qualities of balance, rhythm, symmetry, or unity. Thus the factor analysis was made by "tests" upon "pleasantness" ratings of a 14x50 matrix—composed of the 14 most unevenly weighted "tests" as rated twice by the 50 most reliable subjects.

The Method of the factor analysis of 14 *designs.* The analysis was done by the centroid method, and the axes were rotated in conformance to the criterion of maximizing zero entries in the columns of the factor matrix.

TABLE I.

Matrix of rotated factor loadings
The subscripts of k designate the factor number.

Test (design no.)		k_1	k_2	k_3	k_4	k_5	h^2
1.	(8)	.2297	.5691	.3131	.0807	-.1313	.4984
2.	(13)	.6760	-.0079	.0901	.1563	-.0483	.4919
3.	(15)	.5972	.0183	.0364	-.0371	.1958	.3980
4.	(19)	.2338	-.0947	.0710	.2794	-.0653	.1510
5.	(24)	.1897	-.3281	.3728	.0028	-.2888	.3660
6.	(25)	.1168	-.2016	-.2436	.4078	.2427	.3388
7.	(39)	-.2852	.1501	.5606	.0194	.1032	.4292
8.	(41)	.0166	-.1761	.0062	.4671	-.1985	.2889
9	(42)	.2263	.3093	.1500	-.1528	-.2848	.2738
10.	(43)	.4721	.0929	.0523	.5300	-.2350	.5704
11.	(44)	-.0272	.4017	.1888	.4944	.2484	.5039
12.	(45)	.1153	.5581	-.0109	.0398	.1558	.3508
13.	(48)	.1304	-.0876	.7232	.2913	.0404	.6342
14.	(49)	.5776	-.0897	.2015	.2518	.3565	.5728
	Σ	3.2690	1.1138	2.5114	2.8310	.0907	

Although 5 factors were removed before it could be said that the correlation matrix was completely exhausted, only the first three are discussed in this summary.

The results of factor analyzing the "pleasantness" scores of 14 *designs.* Table I presents the rotated matrix of factor loadings. The designs which have the highest loadings with factor 1 are shown in A of Plate III; those correlating with factor 2 are shown in B; C and D give figures loaded most highly with 3 and 4, respectively. The last group, E, contains the residue of the 14 figures. Factor 5 is not represented.

Inspection of the patterns in A of Plate III suggests that the common characteristic of these is a lack of any unifying or integrating principle among the elements of the design. On the other hand, the parts of the two figures in B seem to have several characteristics which organize them. The designs in C, whose ratings correlate with factor 3, do not fall clearly into either of these categories. There is some slight evidence, secured from verbal reports, that factor 3 may be associated with the connotations of the figures; but the evidence is untrustworthy.

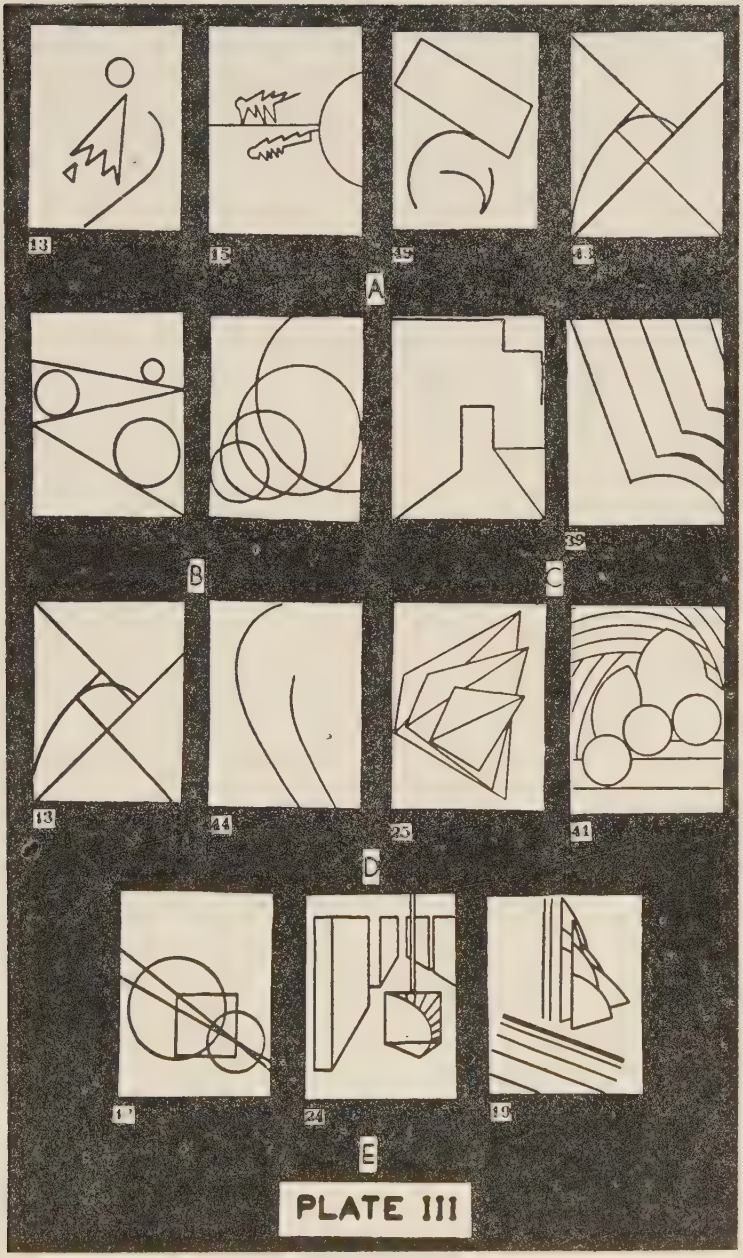

PLATE III

The designs loaded with factor 1 are the lowest of the 14 in "pleasantness" ratings, while those two which correlate with factor 2 are the highest. From this fact, and the observation that the two groups of designs differed in the degree in which they appeared organized, the following hypothesis was made: "pleasantness" is correlated with the ease in which some organizing law can be inferred for the elements of an abstract design. Accordingly, loadings in factor 1 should correlate negatively with "pleasantness", while loadings with factor 2 should correlate positively. Moreover, if evaluations can be secured for "ease of seeing organization", the first two factor loadings should show the same correlations with ratings on this new quality as they did with ratings on "pleasantness".

It was partly to test this hypothesis that a new factor analysis was performed.

The selection of designs for the factor analysis of 10 *figures.* All four designs loaded with factor 1 were retained for this new analysis; one figure loaded with factor 2 was kept; and one which was loaded with all factors. To these 6 were added 4 from the larger group of 50. The four additional designs were chosen because they were given the highest "pleasantness" ratings of the entire battery.

TABLE II.

The rotated factor matrix
The subscripts of k designate the factor number.

Test (design no.)	k_1	k_2	k_3	k_4	k_5	h^2
1. (2)	.1018	.1119	.4912	.2810	-.1244	.3586
2. (4)	.8319	.0949	.0234	.0873	-.2490	.7712
3. (8)	.6182	.1088	.1937	-.0208	.5855	.7748
4. (13)	.0770	.6599	.0978	-.2294	.1568	.5282
5. (15)	.2085	.5993	.0945	-.0209	-.2774	.4890
6. (18)	.3035	-.0153	-.2966	.4647	.1837	.4300
7. (42)	.5331	-.0917	-.2578	-.3485	-.0642	.4846
8. (43)	.2599	.3637	.0415	-.4132	.1861	.4069
9. (49)	.0298	.6289	.0385	-.1720	.0663	.4319
10. (50)	.1880	.0127	.4981	-.1444	-.0053	.3045
Σ	3.1517	2.4731	.9243	-.5162	.4581	

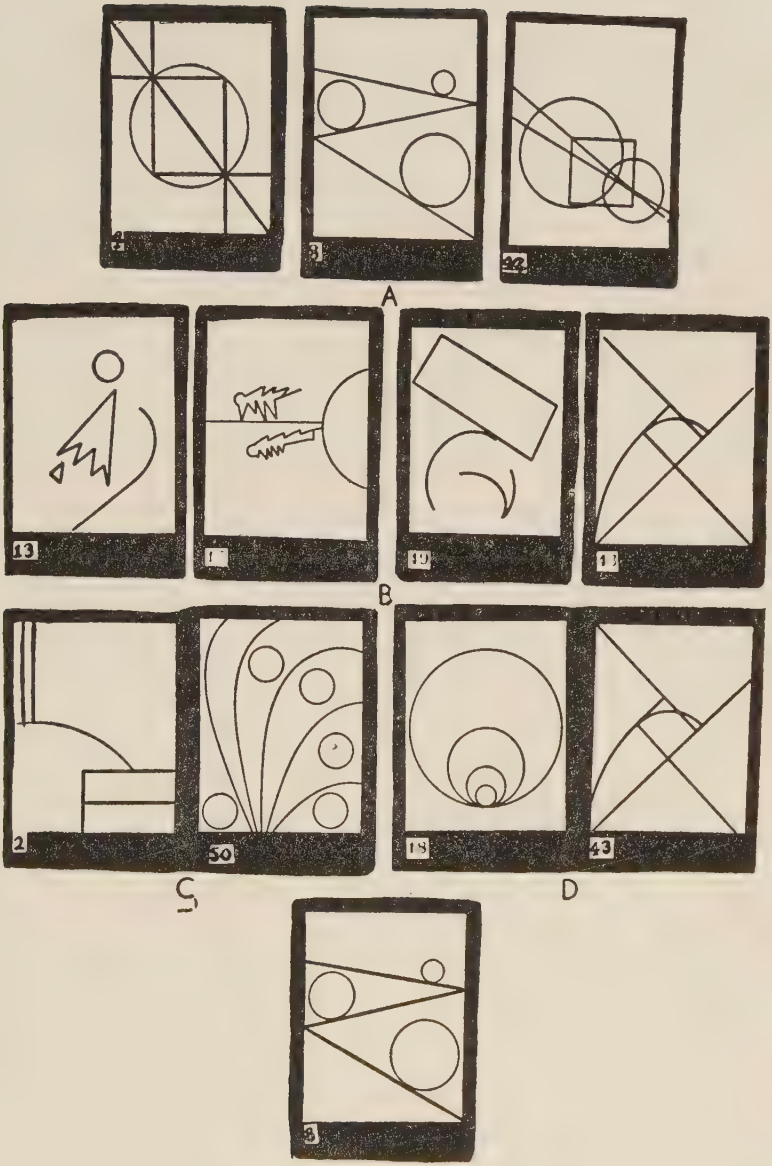

PLATE IV

Method of analysis of the 10 designs. The factor analysis was performed upon the "pleasantness" ratings of the 10 figures just as it was upon the 14. Five factors were removed and rotated by the same criterion. Table II gives the rotated factor loadings which resulted.

Results of the factor analysis of 10 designs. In Plate IV are to be found the designs grouped according to the 5 factors. Their loadings appear in Table II. Factors 1 and 2 correspond to 1 and 2 of the previous analysis, but in reverse order. The other factors do not so agree. Consequently, interpretations of only factors 1 and 2 will be attempted.

Beginning with factor 2, shown in B of Plate IV, it is seen that the entire cluster from the first analysis has been retained intact, and without additions. Thus the presence of the new figures rated high in "pleasantness" leaves this factor unchanged. Factor 1 cannot reappear exactly as it was, however, since one of its designs was omitted from the group of 10. Nevertheless, the retained figure appears in factor 1, and it is accompanied by the design which was loaded with all factors. In addition, a new pattern has been added from the fresh group. All three of these, shown in A of Plate IV, are high in "pleasantness" ratings; hence the factor retains its association with "pleasantness." Likewise, the other factor retains its negative correlation. Moreover, the new factor 1 is again characterized by perceptual organization of the design elements.

The results of testing the hypothesis that "ease of seeing organization" is correlated with "pleasantness." According to the suggestion that "pleasantness" is highly correlated with "ease of seeing organization," factor 1 of the first analysis and factor 2 of the second analysis should correlate negatively with "E.S.O.," and with "pleasantness" ratings. Likewise, the other two corresponding factors of the analyses should correlate positively with "pleasantness" or "E.S.O.," since the figures loaded with them appear to have organization.

The ratings on "E.S.O." were obtained, as described earlier in this section. The factor loadings were then correlated with both these ratings and those on "pleasantness." The coefficients are presented in Table III. Part (a) of the table designates the correlations with loadings from the first analysis, while (b) gives the same for loadings from the second analysis. The subscript of the symbol k designates the factor. Factors 1 and 2 have been permuted in part (b) to emphasize the consistency of the correlations.

TABLE III

Correlations between factor loadings and, (1) ratings in "pleasantness," and (2) "ease of seeing organization." Part (a) contains the results from the 14 design analysis; (b), the 10 figure analysis. The subscript of k indicates the factor.

	k_1	k_2	k_3	k_4	k_5
(a)					
"Pl."	-.689	+.638	-.032	-.148	+.144
"E.S.O."	-.653	+.513	+.058	-.151	-.042
	k_2	k_1	k_3	k_4	k_5
(b)					
"Pl."	-.935	+.388	+.052	+.596	+.116
"E.S.O."	-.933	+.617	+.014	+.543	+.303

The results given in Table III lend support to the hypothesis as it was stated. It was further proved that for the ratings given these particular designs, "E.S.O." and "pleasantness" are about the same thing. This was done by correlating the ratings on these two qualities for the 14 and the 10 designs. For the 14 figures this correlation is +.806; while for the 10 "tests" it is +.958.

The question immediately arises as to whether "E.S.O." can be made to account for "pleasantness" ratings throughout the battery of 50 designs. The summed ratings of each of these qualities were correlated, with the resulting correlation of +.835. When corrected for attenuation, this coefficient became +.875. Thus it seems that the instruction to rate for "ease of seeing organization" yields scores which approximate those obtained under the instruction to rate for "pleasantness"; but it cannot be said that the two instructions produce identical results.

It will be seen in the conclusions of the next experiment that the quality of "ease of seeing organization" can be accounted for in the most part by two "formal criteria"—balance and rhythm.

II "THE ARITSTS' RATING EXPERIMENT"

Subject personnel: The subjects of this experiment were 5 members of the staff of the School of Fine Arts of the University of Illinois. All

were men. On the second evening of the experiment the place of one of the subjects was taken by another member of the same department.

Materials: Four rating scales were substituted for the one on "pleasantness" employed in the previous experiment. These were captioned "symmetry," "balance," "unity," and "rhythm." The segments were headed with degrees of these qualities. Unlike the others, these were scored as 5 point scales only, since one of the subjects reported that he could make no finer discriminations.

Procedure: Instructions were read to the subjects. The first evening they rated on symmetry and balance; the second, on rhythm and unity. While three subjects estimated one quality the other two rated on the other. Their roles were then reversed, thus resulting in ratings upon all four criteria, but in different sequences. No definitions were supplied the subjects. They were told, instead, that a part of the experiment was to learn if agreement exists among artists in the application of these terms.

Results

This experiment is summarized in Tables IV, V, and VI. In Table IV the sums of the ratings by two artists are correlated with the sums of the ratings by the other three. The resulting coefficients of group reliability are given with the same correlations "stepped up" by the application of the prophecy formula for 1.5N. The reliability ranges from +.69 on unity to +.90 on symmetry. If the subject group were comparable in size to the 138 students used in "The Rating Scale Experiment," the reliabilities of mean ratings on "formal criteria" would be higher than the reliability of "pleasantness."

TABLE IV

The reliability coefficients of artists' mean ratings on four undefined qualities of design. Five artists' ratings, "r," obtained by summing 3 vs. 2 subjects, then applying the Spearman-Brown prophecy formula for 1.5N.

Name of quality	r	Prophecy for 1.5N
Symmetry	+.860	+.902
Balance	+.771	+.835
Rhythm	+.816	+.869
Unity	+.699	+.777

In Table V are given the intercorrelations of pooled ratings on balance, symmetry, rhythm, and unity. It can be seen by a comparison of Tables IV and V that ratings in symmetry will predict ratings on balance as well as they will predict other ratings in symmetry. To a less extent all ratings on the "formal criteria" are interrelated.

TABLE V

The intercorrelations of the four "formal criteria of design," summed, as rated by 5 artists

Names of paired qualities	r
Symmetry-balance	+.897
Symmetry-rhythm	+.491
Symmetry-unity	+.870
Balance-rhythm	+.492
Balance-unity	+.832
Rhythm-unity	+.630

In Table VI are the correlations between the summed estimations of the four criteria and factor loadings. Upon comparing Tables III and VI, it is seen that factor loadings from the first analysis correlate about equally with either rhythm or "E.S.O." This led to calculating the correlation between "E.S.O." and rhythm for all 50 designs. It is +.699. The multiple correlation of "E.S.O." with ratings on rhythm and balance is +.791, which approaches the limits of reliability of these latter two criteria, +.77 and +.82, respectively.

TABLE VI

Correlations between factor loadings and ratings in: (1) symmetry, (2) unity, (3) rhythm, as obtained from summed ratings by 5 artists. The factors found in the analysis of 14 designs are correlated in (a); those from the 10 design analysis are in (b). The subscripts of k indicate the factor.

(a)	k_1	k_2	k_3	k_4	k_5
Symmetry	-.206	+.444	-.258	-.132	-.462
Unity	-.358	+.010	-.372	+.192	-.196
Rhythm	-.693	+.604	-.122	-.015	+.044
(b)	k_2	k_1	k_3	k_4	k_5
Symmetry	-.672	+.594	-.354	+.210	-.021
Unity	-.587	+.334	-.173	-.138	+.086
Rhythm	-.721	+.320	-.040	+.273	+.454

Judging from all the evidence presented up to this point, the simplest hypothesis seems to be that the "formal criteria" in some way contribute to "ease of seeing organization," which, in turn, contributes to "pleasantness."

III "THE TRAINING EXPERIMENT"

This study was designed to afford a partial answer to the criticism that there may exist other criteria which account for "pleasantness" more completely than the "formal" ones selected. In addition, it was constructed to discover if ratings on "pleasantness" can be changed by training in rating on the "formal criteria."

Subject personnel: Complete data could be obtained for only 5 persons during the entire six weeks of experimentation. Partial data were available for 13, and some data for as many as 16 were used. The subjects were students, with the exception of one instructor in psychology.

Materials: The subjects received rating scales on "pleasantness," "symmetry," "balance," and "unity," plus three others, labeled, "complexity," "depth," and "meaningfulness." The "depth" and "meaningfulness"

scales were divided into only three segments. Nevertheless, the subjects were informed that all scales were to be regarded as continuous. Each was scored by dividing its segments in half.

Procedure: A complex method of following the act of rating on "pleasantness" by rating on some one of the "formal criteria," then four days later by "pleasantness," made up the general scheme used. Division of the subjects into two groups with different rating sequences, and the use of three individuals as controls, were part of the plan to control transfer of training from one "formal criterion" to another.

Results

The effect of training in "formal criteria" upon "pleasantness" ratings. The correlations used in this experiment were calculated for each subject separately. These were computed between his ratings on "pleasantness" and some one of the "formal criteria" before training and after training. This was done with each of balance, symmetry, rhythm, and unity. If every correlation rose after training in a given rating, the act of rating in that quality was presumed to have had some effect. Data were available for such comparisons to the following extents: 11 pairs in symmetry, 8 in balance, 7 in unity, and 10 in rhythm. All correlations rose after training but those of three persons in the symmetry category only. Other evidence indicates that the definitions given the subjects differed from those used in actual practice in respect to this latter.

The conclusion drawn from these results is that the preference can be changed through training under favorable conditions.

The relationship of "depth," "complexity," and "meaningfulness" to "pleasantness." Coefficients of correlation were obtained between ratings on "complexity" and "pleasantness" for 14 individual subjects. These ranged from -.46 to +.52. While it may be true that "complexity" has a significant negative correlation with "pleasantness" for a few selected persons, it is plain that this quality cannot be made to account for agreement throughout the whole subject group.

If this conclusion is correct for "complexity," it is equally true for "depth." The correlations between "pleasantness" ratings and those for "depth" were distributed normally about a mean of +.12, with a range from -.20 to +.35.

The correlations between "meaningfulness" and "pleasantness" for single persons show greater possibilities of a significant relationship.

Correlations as high as +.50 were obtained for single isolated individuals between their ratings on "meaningfulness" and their initial ratings on "pleasantness." The range was from -.14 to +.53. It was found, however, that only one subject had a higher correlation between these two than he did between "pleasantness" and some one of the "formal criteria." Moreover, a distribution of the intercorrelations between "meaningfulness" ratings of 10 subjtcts was normal about a mean of +.25. The range was from -.06 to +.54. This makes it clear that the subjects failed to agree completely upon what is "meaningful."

It seems, then, that "meaningfulness" cannot be made to account for group agreements upon "pleasantness," since it fails to follow that those few who prefer "meaningful" designs agree with others as to which ones these should be.

SUMMARY

The entire undertaking can be summarized by saying that consistency exists in the aesthetic preferences of untrained individuals for abstract designs. Further, that these consistencies can be accounted for in the most part by a property of the figure which is described as "ease of seeing organization." Moreover, this "ease of seeing organization" may be related to those accepted criteria of design—rhythm, balance, and unity. It has not been shown that the latter are separate factors in the factor analysis sense, or that they exist independently in practical application. Finally, it appears unlikely that considerations of meaning, depth, or complexity play a very influential part in this kind of aesthetic evaluation.

VITA

Llewellyn Noel Wiley was born July 30, 1910, in Vienna, Virginia. Until 1922 he resided in several states at various intervals. From 1922 he has lived in Peoria, Illinois, where he received his high school and undergraduate training.

He received his B.S. degree from Bradley Polytechnic Institute in 1931, and the M.S. degree from Purdue University in 1934.

During the period 1931-'35 he taught in the junior high and high schools of Peoria, Illinois. From this work he moved to an assistantship in psychology in the University of Illinois. He has been with the Department of Psychology from 1935 to 1940.

His memberships include Phi Kappa Phi, Sigma Xi, and three professional societies in psychology. During his enrollment at Bradley he held the Kiwanis Scholarship awarded by the Kiwanis Club of Peoria. He also holds a life teaching certificate in the high schools of Illinois.

In 1936 he published with Prof. Otis C. Trimble on "The Ordinary Objective Test as a Possible Criterion of Certain Personality Traits"; and in 1939 on the validity of formal criteria in abstract design.